

РАСПОЗНАВАНИЕ РУКОПИСНЫХ ЦИФР НА ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ ИНСТРУМЕНТОВ МАШИННОГО ОБУЧЕНИЯ

А.С. Вторушина, И.А. Ботыгин
Томский политехнический университет
E-mail: chibi.moon@mail.ru

Введение

Среди существующих задач современной кибернетики, стоит отметить проблему распознавания образов рукописного ввода цифровых символов. В настоящее время данная задача во многом популярна благодаря пристальному вниманию многих исследователей. В данной работе рассматриваются принципы и методы классификации и идентификации объектов, явлений, процессов, сигналов и ситуаций.

Среди основных подходов к решению задач распознавания объектов стоит выделить: классификации с помощью функций, различные алгоритмы кластеризации, нейронные сети и т.д. Главная цель распознавания образов основана на обучающей выборке для определения класса, к которому необходимо классифицировать или идентифицировать описание объекта. Любое решение проблемы предусматривается для этой схемы, если процесс такого принятия в первую очередь основан на изучении ранее приобретенного опыта с задачами распознавания образов, а также включает задачу минимизации описания исходных объектов и извлечения информационных функций [1]. Особый интерес представляют искусственные нейронные сети, строящиеся на алгоритмах обучения, где в качестве тестовой выборки выступают открытые базы различных изображений.

Результаты эксперимента

Целью экспериментальной части работы является обеспечение максимальной точности распознавания нейронной сетью рукописных числовых знаков в диапазоне от 0 до 9. При этом в процессе работы нейросетевого алгоритма нагрузка на аппаратную часть вычислительной машины должна быть адекватной и равномерной [2-4]. К аппаратной части машины в данном случае стоит отнести центральный и графический процессоры.

В процессе эксперимента основными задачами ставились исследование влияния на точность выходных данных таких параметров как:

1. количество слоев сверточной нейросети;
2. количество циклов обучения (эпох);
3. количество сетов эпохи (Batch size);
4. соотношение тренировочных и обучающих объектов обучающего множества (Validation split).

При построении модели нейронной сети следует обратить внимание на одну из основных проблем – переобучение. Переобучение возникает в случае слишком долгого обучения, недостаточного числа обучающих примеров или переусложненной структуры нейронной сети.

Первым этапом эксперимента является определение зависимости ошибки обучения от параметра Validation split. Варьирование данного параметра служит одним из вариантов борьбы с переобучением сети. Также он отвечает за деление обучающего множества на два множества – обучающее и тестовое.

Количество эпох и слоев сети следует выбрать минимальным для исключения влияния этого параметра на функцию ошибок. Количество сетов эпохи следует выбрать исходя из возможностей аппаратной части вычислительной машины. Для примера, достаточно выбрать 2 параметра Batch size: 100 и 200. Диапазон варьирования параметра Validation split принят в пределах от 0,1 до 0,9. Графики приведен ниже (рис. 1).

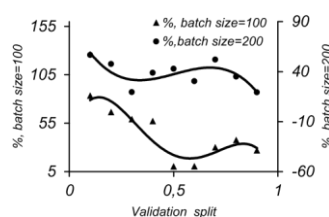


Рис. 1. Точность

При разных количествах сетов в эпохе можно заметить снижение точности выходных данных при Validation split, стремящемуся к 1. Это говорит о том, что чем больше объектов обучения общего множества выделено на обучение сети, тем точность обучения выше. Функция ошибок Loss ведет себя несколько иначе. При малых значениях Batch size четкой корреляции и тенденции не прослеживается, в то время как при увеличении Batch size заметно увеличение точности выходных данных при соотношении обучающего и тестового множеств 50/50. При других соотношениях точность обучения имеет тенденцию к снижению.

Построенная нейронная сеть, основанная на 1 сверточном слое, показывает недостаточную точность выходных данных при оптимальном количестве сетов одной эпохи во всем диапазоне Validation split. Для сравнения по точности достаточно внедрить в сеть еще один слой и выбрать оптимальные параметры сети (2).

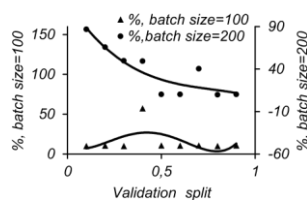


Рис. 2. Точность обучения в процентном соотношении

По аналогии с предыдущей однослойной сетью более выраженный спад точности имеет нейронная сеть с большим количеством Batch size.

В сравнении с однослойной нейросетью, в двухслойной наблюдается наиболее ярко выраженные тенденции кривых. В то же время одни и те же показатели качества обеих сетей коррелируют между собой. К примеру, показатель точности сети, содержащей один сверточный слой, имеет тенденцию к снижению качественной характеристики практически независимо от количества сетов в одной эпохе. Этот же показатель точности имеет более выраженную тенденцию к спаду в нейронной сети с двумя слоями. При этом наибольшую крутизну кривой имеет конфигурация сети с большим числом сетов. Величина Batch size в значительной степени влияет на производительность сети и загруженность локальной машины. Конфигурация с Batch size = 200 для данной машины вполне приемлема, однако, такая нейросеть не обладает достаточной точностью выходных данных. Поэтому следует повысить количество эпох вплоть до десяти, зафиксировав при этом величину Batch size в размере двухсот объектов.

Наибольшей точностью обладает сеть с максимальным количеством (десятью) сверточных слоев (рис.3).

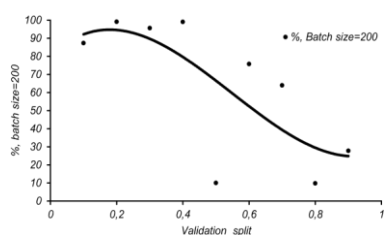


Рис. 3. Точность

В зависимости от соотношения обучающего и тестового множества точность выходных данных имеет тенденцию, преимущественно, к падению. Однако максимальная точность достигается при соотношении 20% – тестовых данных и 80% – обучающих.

Результат распознавания рукописных чисел представлены на рисунке 4.

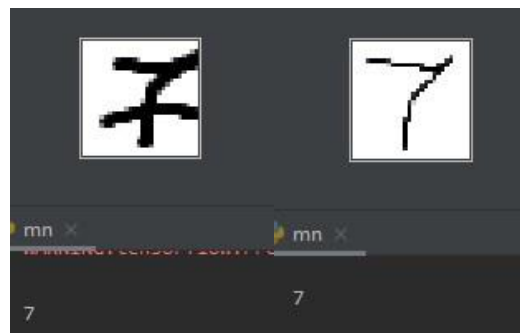


Рис. 4. Распознавание цифры семь.

Заключение

В результате проделанной работы было проведено исследование принципа работы сверточной нейронной сети и ее прикладное применение для распознавания образов, в частности – рукописных цифр при помощи современных методов машинного обучения. Разработано консольное приложение на языке программирования Python для работы с данными из открытой базы MNIST. Для MNIST задействуется сверточная нейросеть с 6 слоями.

Разработанное приложение впоследствии может послужить отправной точкой для разработки ядра более мощной программы распознавания рукописного ввода, а именно связки цифр, символов (номера автомобилей, почтовые индексы и т.д.).

Список использованных источников

1. Тадеева Н.В. Нейронное распознавание образов// Национальная ассоциация ученых. – 2015 –№2-4(5) –С.71-74. [Электронный ресурс]. – Режим доступа: <https://elibrary.ru/item.asp?id=29111296>. Дата обращения: 22 февраля 2019.
2. Сверточные нейронные сети. Интернет-энциклопедия. [Электронный ресурс]. – Режим доступа: http://neerc.ifmo.ru/wiki/index.php?title=%D0%A1%D0%B2%D0%B5%D1%80%D1%82%D0%B5%D1%87%D0%BD%D1%8B%D0%B5_%D0%BD%D0%B5%D0%B9%D1%80%D0%BE%D0%BD%D0%BD%D1%8B%D0%B5_%D1%81%D0%B5%D1%82%D0%B8. Дата обращения: 3 марта 2019.
3. Попов Н. А. Анализ потока людей средствами компьютерного зрения: бакалаврская работа / Н. А. Попов; Национальный исследовательский Томский политехнический университет (ТПУ), Инженерная школа ядерных технологий (ИЯТШ), Отделение экспериментальной физики (ОЭФ); науч. рук. М. Е. Семенов. — Томск, 2018.
4. НАВР. Интернет-портал для IT-специалистов. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/348000>. Дата обращения: 3 мая 2019.